PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/47, C12N 5/10, C07K 16/18, A61K 38/17, C12Q 1/68

A1

(11) International Publication Number: **WO 00/18919**

(43) International Publication Date: 6 April 2000 (06.04.00)

(21) International Application Number: PCT/US99/22495

(22) International Filing Date: 28 September 1999 (28.09.99)

(30) Priority Data:
09/162,373 28 September 1998 (28.09.98) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 09/162,373 (CIP)
Filed on 28 September 1998 (28.09.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way, #1, Menlo Park, CA 94025 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HUMAN KINESIN–LIKE MOTOR PROTEIN

(57) Abstract

The invention provides a human kinesin–like motor protein (KLIMP) and polynucleotides which identify and encode KLIMP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of KLIMP.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

KINESIN-LIKE MOTOR PROTEIN

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of a kinesin-like motor protein and to the use of these sequences in the diagnosis, treatment, and prevention of cancer, neurological disorders, and disorders of vesicular transport.

BACKGROUND OF THE INVENTION

Translocation of components within the cell is critical for maintaining cell structure and function. Cellular components such as proteins and membrane-bound organelles are transported along well-defined routes to specific subcellular compartments. Intracellular transport mechanisms utilize microtubules which are filamentous polymers that serve as tracks for directing the movement of molecules. Molecular transport is driven by the microtubule-based motor proteins, kinesin and dynein. These proteins use the energy derived from ATP hydrolysis to power their movement unidirectionally along microtubules and to transport molecular cargo to specific destinations.

Kinesin defines a ubiquitous, conserved family of over 50 proteins that can be classified into at least 8 subfamilies based on primary amino acid sequence, domain structure, velocity of movement, and cellular function. (Reviewed in Moore, J.D. and S.A. Endow (1996) Bioessays 18:207-219; Hoyt, A.M. (1994) Curr. Opin. Cell Biol. 6:63-68.) The prototypical kinesin molecule is involved in the transport of membrane-bound vesicles and organelles. This function is particularly important for axonal transport in neurons. Protein-containing vesicles are constantly transported from the neuronal cell body along microtubules that span the length of the axon leading to the synaptic terminal. Failure to supply the synaptic terminal with these vesicles blocks the transmission of neural signals. In the fruit fly Drosophila melanogaster, for example, mutations in kinesin cause severe disruption of axonal transport in larval nerves which leads to progressive paralysis (Hurd, D.D. and W.M. Saxton (1996) Genetics 144:1075-1085). This phenotype mimics the pathology of some vertebrate motor neuron diseases, such as amyotrophic lateral sclerosis (ALS). In addition to axonal transport, kinesin is also important in all cell types for the transport of vesicles from the Golgi complex to the endoplasmic reticulum. This role is critical for maintaining the identity and functionality of these secretory organelles.

Members of the more divergent subfamilies of kinesin are called kinesin-related proteins (KRPs), many of which function during mitosis in eukaryotes as divergent as yeast and human (Hoyt, supra). Some KRPs are required for assembly of the mitotic spindle. In vivo and in vitro

analyses suggest that these KRPs exert force on microtubules that comprise the mitotic spindle, resulting in the separation of spindle poles. Phosphorylation of KRP is required for this activity. Failure to assemble the mitotic spindle results in abortive mitosis and chromosomal aneuploidy, the latter condition being characteristic of cancer cells. In addition, a unique KRP, centromere protein E, localizes to the kinetochore of human mitotic chromosomes and may play a role in their segregation to opposite spindle poles.

The prototypical kinesin molecule is a heterotetramer comprised of two heavy polypeptide chains (KHCs) and two light polypeptide chains (KLCs). The KHC subunits are typically referred to as "kinesin." KHC is about 1000 amino acids in length, and KLC is about 550 amino acids in length. Two KHCs dimerize to form a rod-shaped molecule with three distinct regions of secondary structure. At one end of the molecule is a globular motor domain that functions in ATP hydrolysis and microtubule binding. Kinesin motor domains are highly conserved and share over 70% identity. Beyond the motor domain is an α-helical coiled-coil region which mediates dimerization. At the other end of the molecule is a fan-shaped tail that associates with molecular cargo. The tail is formed by the interaction of the KHC C-termini with the two KLCs.

The nematode Unc-104 kinesin-like protein defines a distinct kinesin subfamily whose members may function monomerically (Moore and Endow, supra). Members of this subfamily are important for synaptic transport and mitochondrial translocation and are characterized by movement at relatively high velocities of about 40 to 90 microns/minute. Nematodes with mutations in the Unc-104 gene exhibit defects in locomotion and feeding behaviors and, at the molecular level, in synaptic vesicle transport.

The discovery of a new kinesin-like motor protein and the polynucleotides encoding it satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cancer, neurological disorders, and disorders of vesicular transport.

SUMMARY OF THE INVENTION

The invention is based on the discovery of a new human kinesin-like motor protein (KLIMP), the polynucleotides encoding KLIMP, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, neurological disorders, and disorders of vesicular transport.

The invention features a substantially purified polypeptide comprising the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1.

The invention further provides a substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID

NO:1. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1. The invention also includes an isolated and purified polynucleotide variant having at least 80% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1.

The invention further provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1, as well as an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1.

The invention also provides an isolated and purified polynucleotide comprising the polynucleotide sequence of SEQ ID NO:2 or a fragment of SEQ ID NO:2, and an isolated and purified polynucleotide variant having at least 80% polynucleotide sequence identity to the polynucleotide comprising the polynucleotide sequence of SEQ ID NO:2 or a fragment of SEQ ID NO:2. The invention also provides an isolated and purified polynucleotide having a sequence complementary to the polynucleotide comprising the polynucleotide sequence of SEQ ID NO:2 or a fragment of SEQ ID NO:2.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising the sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1 in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide

comprising the sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1, as well as a purified agonist and a purified antagonist of the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of KLIMP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of KLIMP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of the polypeptide having the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1.

BRIEF DESCRIPTION OF THE FIGURES AND TABLE

Figures 1A, 1B, 1C, 1D, 1E, 1F, 1G, 1H, 1I, and 1J show the amino acid sequence (SEQ ID NO:1) and nucleic acid sequence (SEQ ID NO:2) of KLIMP. The alignment was produced using MACDNASIS PRO software (Hitachi Software Engineering, S. San Francisco CA).

Figure 2 shows electronic northern analysis of SEQ ID NO:2 using the LIFESEQ sequence database (Incyte Pharmaceuticals, Palo Alto CA).

Table 1 shows the programs, their descriptions, references, and threshold parameters used to analyze KLIMP.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention

belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"KLIMP" refers to the amino acid sequences of substantially purified KLIMP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to KLIMP, increases or prolongs the duration of the effect of KLIMP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of KLIMP.

An "allelic variant" is an alternative form of the gene encoding KLIMP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding KLIMP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as KLIMP or a polypeptide with at least one functional characteristic of KLIMP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding KLIMP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding KLIMP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent KLIMP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of KLIMP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic

acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of KLIMP which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of KLIMP. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to KLIMP, decreases the amount or the duration of the effect of the biological or immunological activity of KLIMP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of KLIMP.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind KLIMP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic KLIMP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding KLIMP or fragments of KLIMP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW Fragment Assembly system (GCG, Madison WI). Some sequences have been both

extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding KLIMP, by northern analysis is indicative of the presence of nucleic acids encoding KLIMP in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding KLIMP.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988)

Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" or "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" or "array element" in a microarray context, refer to hybridizable

polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of KLIMP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of KLIMP.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:2, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:2 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:2 from related polynucleotide sequences. A fragment of SEQ ID NO:2 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:2 and the region of SEQ ID NO:2 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" or "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript

elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding KLIMP, or fragments thereof, or KLIMP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment.

The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of KLIMP polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to KLIMP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of a new human kinesin-like motor protein (KLIMP), the polynucleotides encoding KLIMP, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, neurological disorders, and disorders of vesicular transport.

Nucleic acids encoding the KLIMP of the present invention were identified in Incyte Clone 1281811 from the colon cDNA library (COLNNOT16) using a computer search for nucleotide and/or amino acid sequence alignments. A consensus sequence, SEQ ID NO:2, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones

1281827H1 and 1281811H1 (COLNNOT16), 3098665H1 (CERVNOT03), 3402610H1 (ESOGNOT03), 893899X28F1 and 964318X27 (BRSTNOT05), 3386436H1 (LUNGTUT17), 660598X15 (BRAINOT03), 2791542F6 and 2791542T6 (COLNTUT16), 3046745H1 (HEAANOT01), 1988294R6 (LUNGAST01), and 1257207F1 (MENITUT03).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1, as shown in Figures 1A, 1B, 1C, 1D, 1E, 1F, 1G, 1H, 1I, and 1J. KLIMP is 1103 amino acids in length and has six potential N-glycosylation sites at N33, N133, N211, N272, N455, and N851; two potential cAMP- and cGMP-dependent protein kinase phosphorylation sites at S295 and S33; eighteen potential casein kinase II phosphorylation sites at T20, T60, S61, S145, S188, T234, S245, S326, S385, T457, S676, S684, S694, T743, T806, S852, S919, and S1092; nineteen potential protein kinase C phosphorylation sites at S5, S165, S214, T226, S238, S252, S290, T338, S396, T447, S494, S680, S688, S694, S716, S852, S905, S910, and S1022; and one potential tyrosine kinase phosphorylation site at Y300. MOTIFS analysis shows that KLIMP contains an ATP-binding motif from G97 to S104. PROFILESCAN and PFAM indicate the presence of a kinesin motor domain from R11 to L377. Within this region, MOTIFS analysis identifies a kinesin motor domain signature sequence from S242 to E253; BLOCKS analysis indicates that KLIMP contains five out of eight protein domain blocks which are characteristic of kinesin motor domains and which are most closely related to those blocks found in Unc-104; and PRINTS analysis indicates that KLIMP contains four out of four protein fingerprints which are likewise characteristic of kinesin motor domains. A fragment of SEQ ID NO:2 from about nucleotide 1518 to about nucleotide 1547 is useful in hybridization or amplification technologies to identify SEQ ID NO:2 and to distinguish between SEQ ID NO:2 and a related sequence.

Northern analysis shows the expression of this sequence in various libraries, at least 68% of which are associated with cancer or cell proliferation. In particular, 27% of the libraries expressing KLIMP are derived from reproductive tissue and 19% are derived from neural tissue. Figure 2 shows the three cDNA libraries from the LIFESEQ database in which SEQ ID NO:2 is most abundantly expressed. Abundance refers to the number of times SEQ ID NO:2 appears in each of the libraries listed, and percent abundance refers to the abundance divided by the total number of sequences examined in a given library. Of particular note is that these three cDNA libraries are all derived from neurodegenerative tissue: MUSCNOT11 is derived from diseased muscle tissue from a patient who died of amyotrophic lateral sclerosis, BRAYDIT01 is derived from hypothalamus tissue from a patient with Huntington's disease, and PONSAZT01 is derived from diseased pons tissue removed from the brain of a patient who died of Alzheimer's disease.

The invention also encompasses KLIMP variants. A preferred KLIMP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the KLIMP amino acid sequence, and which contains at least one functional or structural characteristic of KLIMP.

The invention also encompasses polynucleotides which encode KLIMP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising the sequence of SEQ ID NO:2, which encodes KLIMP.

The invention also encompasses a variant of a polynucleotide sequence encoding KLIMP. In particular, such a variant polynucleotide sequence will have at least about 80%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding KLIMP. A particular aspect of the invention encompasses a variant of SEQ ID NO:2 which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to SEQ ID NO:2. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of KLIMP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding KLIMP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring KLIMP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode KLIMP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring KLIMP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding KLIMP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding KLIMP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode KLIMP and KLIMP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding KLIMP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:2, or to a fragment of SEQ ID NO:2, under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 µg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 µg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3

mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the Hydra microdispenser (Robbins Scientific, Sunnyvale CA), MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (Perkin-Elmer). Sequencing is then carried out using either ABI 373 or 377 DNA sequencing systems (Perkin-Elmer) or the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA). The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding KLIMP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR,

nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode KLIMP may be cloned in recombinant DNA molecules that direct expression of KLIMP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express KLIMP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter KLIMP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding KLIMP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, KLIMP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of KLIMP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active KLIMP, the nucleotide sequences encoding KLIMP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding KLIMP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding KLIMP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding KLIMP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding KLIMP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques,

synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding KLIMP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding KLIMP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding KLIMP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding KLIMP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of KLIMP are needed, e.g. for the production of antibodies, vectors which direct high level expression of KLIMP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of KLIMP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of KLIMP. Transcription of sequences encoding KLIMP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV

used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding KLIMP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses KLIMP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of KLIMP in cell lines is preferred. For example, sequences encoding KLIMP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ or *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers

resistance to methotrexate; *neo* confers resistance to the aminoglycosides, neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding KLIMP is inserted within a marker gene sequence, transformed cells containing sequences encoding KLIMP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding KLIMP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding KLIMP and that express KLIMP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of KLIMP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on KLIMP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding KLIMP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding KLIMP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding KLIMP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode KLIMP may be designed to contain signal sequences which direct secretion of KLIMP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding KLIMP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric KLIMP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of KLIMP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using

commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the KLIMP encoding sequence and the heterologous protein sequence, so that KLIMP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled KLIMP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of KLIMP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of KLIMP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of KLIMP and the motor domain of kinesin. In addition, the expression of KLIMP is closely associated with cancerous and proliferating cells and neurodegenerative tissue. Therefore, KLIMP appears to be associated with cancer, neurological disorders, and disorders of vesicular transport. In the treatment of cancer, neurological disorders, and disorders of vesicular transport associated with increased KLIMP activity, it is desirable to decrease the expression or activity of KLIMP. In the treatment of the above conditions associated with decreased KLIMP activity, it is desirable to provide the protein or to increase the expression of KLIMP.

Therefore, in one embodiment, KLIMP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or

activity of KLIMP. Examples of such disorders include, but are not limited to, cancers such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; neurological disorders such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; and disorders of vesicular transport such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, gastrointestinal disorders including ulcerative colitis, gastric and duodenal ulcers, other conditions associated with abnormal vesicle trafficking including acquired immunodeficiency syndrome (AIDS), allergic reactions, autoimmune hemolytic anemia, proliferative glomerulonephritis, inflammatory bowel disease, multiple sclerosis, myasthenia gravis, rheumatoid arthritis, osteoarthritis, scleroderma, Chediak-Higashi syndrome, Sjogren's syndrome, systemic lupus erythematosus, toxic shock syndrome, traumatic tissue damage, and viral, bacterial, fungal, helminthic, and protozoal infections.

In another embodiment, a vector capable of expressing KLIMP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased

expression or activity of KLIMP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified KLIMP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of KLIMP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of KLIMP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of KLIMP including, but not limited to, those listed above.

In a further embodiment, an antagonist of KLIMP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of KLIMP. Such disorders may include, but are not limited to, those discussed above. In one aspect, an antibody which specifically binds KLIMP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express KLIMP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding KLIMP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of KLIMP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of KLIMP may be produced using methods which are generally known in the art. In particular, purified KLIMP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind KLIMP. Antibodies to KLIMP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with KLIMP or with any fragment or

oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to KLIMP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of KLIMP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to KLIMP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce KLIMP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for KLIMP may also be

generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between KLIMP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering KLIMP epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for KLIMP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of KLIMP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple KLIMP epitopes, represents the average affinity, or avidity, of the antibodies for KLIMP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular KLIMP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} l/mole are preferred for use in immunoassays in which the KLIMP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 l/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of KLIMP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of KLIMP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding KLIMP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding KLIMP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding KLIMP. Thus, complementary molecules or fragments may be used to modulate KLIMP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding KLIMP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding KLIMP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding KLIMP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding KLIMP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding KLIMP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the

ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding KLIMP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding KLIMP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits,

monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of KLIMP, antibodies to KLIMP, and mimetics, agonists, antagonists, or inhibitors of KLIMP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated

sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of

KLIMP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example KLIMP or fragments thereof, antibodies of KLIMP, and agonists, antagonists or inhibitors of KLIMP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of therapeutic to toxic effects is the therapeutic index, and it can be expressed as the ED_{50}/LD_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their

inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind KLIMP may be used for the diagnosis of disorders characterized by expression of KLIMP, or in assays to monitor patients being treated with KLIMP or agonists, antagonists, or inhibitors of KLIMP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for KLIMP include methods which utilize the antibody and a label to detect KLIMP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring KLIMP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of KLIMP expression. Normal or standard values for KLIMP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to KLIMP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of KLIMP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding KLIMP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of KLIMP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of KLIMP, and to monitor regulation of KLIMP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding KLIMP or closely related molecules may be used to identify nucleic acid sequences which encode KLIMP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies

only naturally occurring sequences encoding KLIMP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the KLIMP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:2 or from genomic sequences including promoters, enhancers, and introns of the KLIMP gene.

Means for producing specific hybridization probes for DNAs encoding KLIMP include the cloning of polynucleotide sequences encoding KLIMP or KLIMP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding KLIMP may be used for the diagnosis of disorders associated with expression of KLIMP. Examples of such disorders include, but are not limited to, cancers such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; neurological disorders such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies,

myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; and disorders of vesicular transport such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, gastrointestinal disorders including ulcerative colitis, gastric and duodenal ulcers, other conditions associated with abnormal vesicle trafficking including acquired immunodeficiency syndrome (AIDS), allergic reactions, autoimmune hemolytic anemia, proliferative glomerulonephritis, inflammatory bowel disease, multiple sclerosis, myasthenia gravis, rheumatoid arthritis, osteoarthritis, scleroderma, Chediak-Higashi syndrome, Sjogren's syndrome, systemic lupus erythematosus, toxic shock syndrome, traumatic tissue damage, and viral, bacterial, fungal, helminthic, and protozoal infections. The polynucleotide sequences encoding KLIMP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered KLIMP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding KLIMP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding KLIMP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding KLIMP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of KLIMP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding KLIMP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with

values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or over-expressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding KLIMP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding KLIMP, or a fragment of a polynucleotide complementary to the polynucleotide encoding KLIMP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of KLIMP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See,

e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding KLIMP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding KLIMP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, KLIMP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of

drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between KLIMP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with KLIMP, or fragments thereof, and washed. Bound KLIMP is then detected by methods well known in the art. Purified KLIMP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding KLIMP specifically compete with a test compound for binding KLIMP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with KLIMP.

In additional embodiments, the nucleotide sequences which encode KLIMP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 09/162,373, are hereby expressly incorporated by reference.

EXAMPLES

I. cDNA Library Construction

The COLNNOT16 library was constructed using RNA isolated from sigmoid colon tissue removed from a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy. Pathology for the associated tumor tissue indicated invasive grade 2 adenocarcinoma. One lymph node contained metastasis with extranodal extension. Patient history included hyperlipidemia, cataract disorder, dermatitis, cholecystectomy, and inguinal hernia repair. Family history included

benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, breast cancer, and prostate cancer.

Frozen tissue was homogenized and lysed in guanidinium isothiocyanate solution using a Polytron PT-3000 homogenizer (Brinkmann Instruments, Westbury NY). The lysate was centrifuged over a CsCl cushion to isolate RNA. The RNA was extracted with acid phenol, precipitated with sodium acetate and ethanol, resuspended in RNase-free water, and treated with DNase. The RNA was re-extracted with acid phenol and reprecipitated with sodium acetate and ethanol. Poly(A+) RNA was isolated using the OLIGOTEX mRNA purification kit (QIAGEN, Chatsworth CA).

Poly(A+) RNA was used for cDNA synthesis and construction of the cDNA library according to the recommended protocols in the SUPERSCRIPT plasmid system (Life Technologies). The cDNAs were fractionated on a SEPHAROSE CL4B column (Amersham Pharmacia Biotech), and those cDNAs exceeding 400 bp were ligated into pINCY (Incyte Pharmaceuticals). Recombinant plasmids were transformed into DH5α competent cells (Life Technologies).

II. Isolation of cDNA Clones

Plasmid DNA was released from the cells and purified using the R.E.A.L. Prep 96 plasmid kit (QIAGEN). The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/l and glycerol at 0.4%; 2) after the cultures were incubated for 19 hours, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellets were each resuspended in 0.1 ml of distilled water. The DNA samples were stored at 4°C.

III. Sequencing and Analysis

The cDNAs were prepared for sequencing using the ABI CATALYST 800 (Perkin-Elmer) or the HYDRA microdispenser (Robbins Scientific) or MICROLAB 2200 (Hamilton) systems in combination with the PTC-200 thermal cyclers (MJ Research). The cDNAs were sequenced using the ABI PRISM 373 or 377 sequencing systems (Perkin-Elmer) and standard ABI protocols, base calling software, and kits. In one alternative, cDNAs were sequenced using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics). In another alternative, the cDNAs were amplified and sequenced using the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). In yet another alternative, cDNAs were sequenced using solutions and dyes from Amersham Pharmacia Biotech. Reading frames for the ESTs were determined using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA, extension, and shotgun sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 1 summarizes the software programs, descriptions, references, and threshold parameters used. The first column of Table 1 shows the tools, programs, and algorithms used, the second column provides a brief description thereof, the third column presents the references which are incorporated by reference herein, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the probability the greater the homology). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering) and LASERGENE software (DNASTAR).

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were used to identify polynucleotide sequence fragments from SEQ ID NO:2. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ database (Incyte Pharmaceuticals, Palo Alto CA). This analysis is much faster than multiple membrane-based

hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{\% sequence identity x \% maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported a percentage distribution of libraries in which the transcript encoding KLIMP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease categories included cancer, inflammation/trauma, fetal, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease expression are reported in the description of the invention.

V. Extension of KLIMP Encoding Polynucleotides

The full length nucleic acid sequence of SEQ ID NO:2 was produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the

following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICO GREEN quantitation reagent (0.25% (v/v) PICO GREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with *Pfu* DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequence of SEQ ID NO:2 is used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an

appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 µCi of [^{32}P]-adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT-AR film (Eastman Kodak, Rochester NY) is exposed to the blots for several hours, hybridization patterns are compared visually.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are

arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the KLIMP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring KLIMP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of KLIMP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the KLIMP-encoding transcript.

IX. Expression of KLIMP

Expression and purification of KLIMP are achieved using bacterial or virus-based expression systems. For expression of KLIMP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express KLIMP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of KLIMP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding KLIMP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, KLIMP is synthesized as a fusion protein with, e.g.,

glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from KLIMP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified KLIMP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of KLIMP Activity

A microtubule motility assay for KLIMP activity measures motor domain function. In this assay, recombinant KLIMP is immobilized onto a glass slide or similar substrate. Taxol-stabilized bovine brain microtubules (commercially available) in a solution containing ATP and cytosolic extract are perfused onto the slide. Movement of microtubules as driven by KLIMP motor activity can be visualized and quantified using video-enhanced light microscopy and image analysis techniques. KLIMP activity is directly proportional to the frequency and velocity of microtubule movement.

XI. Functional Assays

KLIMP function is assessed by expressing the sequences encoding KLIMP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP, and to evaluate cellular properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that

diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of KLIMP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding KLIMP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding KLIMP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of KLIMP Specific Antibodies

KLIMP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the KLIMP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide Synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring KLIMP Using Specific Antibodies

Naturally occurring or recombinant KLIMP is substantially purified by immunoaffinity

chromatography using antibodies specific for KLIMP. An immunoaffinity column is constructed by covalently coupling anti-KLIMP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing KLIMP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of KLIMP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/KLIMP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and KLIMP is collected.

XIV. Identification of Molecules Which Interact with KLIMP

KLIMP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled KLIMP, washed, and any wells with labeled KLIMP complex are assayed. Data obtained using different concentrations of KLIMP are used to calculate values for the number, affinity, and association of KLIMP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less
PFAM	A Hidden Markov Models-based application useful for protein family search.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits, depending on individual protein families

Table 1 cont.

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Score= 4.0 or greater
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising the amino acid sequence of SEQ ID NO:1 or a fragment thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 80% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising the polynucleotide sequence of SEQ ID NO:2 or a fragment thereof.

10. An isolated and purified polynucleotide variant having at least 80% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

(a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

(b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of KLIMP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of KLIMP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

```
                11                  20                  29                  38                  47                  56
5'GTGGC AGC CAG AAC TGA TAC AGC CCC CCT GGT CTG GGG CCA GGA CGC CAG CTG AGG

                65                  74                  83                  92                 101                 110
    AGG GCA GGA GTG TCT GGA GCT ATG GCT GGT GCC TCG GTG AAA GTG GCA GTG AGG
                                    M   A   G   A   S   V   K   V   A   V   R

               119                 128                 137                 146                 155                 164
    GTT CGG CCC TTT AAC GCC CGT GAG ACC AGC CAG GAT GCC AAG TGT GTG GTC AGC
    V   R   P   F   N   A   R   E   T   S   Q   D   A   K   C   V   V   S

               173                 182                 191                 200                 209                 218
    ATG CAG GGC AAC ACC ACC TCC ATC ATC AAT CCT AAA CAG AGC AAG GAT GCC CCC
    M   Q   G   N   T   T   S   I   I   N   P   K   Q   S   K   D   A   P

               227                 236                 245                 254                 263                 272
    AAA AGC TTC ACC TTT GAC TAC TCC TAC TGG TCA CAC ACT TCG ACG GAG GAC CCC
    K   S   F   T   F   D   Y   S   Y   W   S   H   T   S   T   E   D   P

               281                 290                 299                 308                 317                 326
    CAG TTT GCA TCT CAG CAG CAA GTG TAT CGG GAC ATT GGA GAA GAG ATG CTG CTC
    Q   F   A   S   Q   Q   Q   V   Y   R   D   I   G   E   E   M   L   L

               335                 344                 353                 362                 371                 380
    CAC GCC TTT GAA GGC TAC AAC GTG TGC ATC TTT GCC TAT GGG CAG ACC GGG GCT
    H   A   F   E   G   Y   N   V   C   I   F   A   Y   G   Q   T   G   A
```

FIGURE 1A

```
          389         398         407         416         425         434
GGG AAA TCC TAT ACC ATG ATG GGG CGA CAG GAG CCA GGG CAG CAG GGC ATC GTG
G   K   S   Y   T   M   M   G   R   Q   E   P   G   Q   Q   G   I   V

          443         452         461         470         479         488
CCC CAG CTC TGT GAG GAC CTC TTC TCT CGC GTT AGT GAG AAC CAG AGT GCT CAG
P   Q   L   C   E   D   L   F   S   R   V   S   E   N   Q   S   A   Q

          497         506         515         524         533         542
CTA TCC TAC TCT GTG GAG GTG AGC TAT ATG GAG ATC TAC TGT GAG CGG GTA CGA
L   S   Y   S   V   E   V   S   Y   M   E   I   Y   C   E   R   V   R

          551         560         569         578         587         596
GAC CTC TTG AAC CCC AAG AGT CGG GGT TCT CTG CGG GTC CGG GAG CAC CCC ATC
D   L   L   N   P   K   S   R   G   S   L   R   V   R   E   H   P   I

          605         614         623         632         641         650
CTG GGC CCG TAC GTG CAG GAC CTG TCC AAA TTG GCT GTG ACC TCC TAC GCA GAC
L   G   P   Y   V   Q   D   L   S   K   L   A   V   T   S   Y   A   D

          659         668         677         686         695         704
ATT GCT GAC CTC ATG GAC TGT GGA AAT AAA GCA CGG ACT GTG GCT GCC ACC AAC
I   A   D   L   M   D   C   G   N   K   A   R   T   V   A   A   T   N

          713         722         731         740         749         758
ATG AAT GAG ACC AGC AGC CGT TCC CAT GCC GTC TTT ACC ATC GTC TTC ACA CAG
M   N   E   T   S   S   R   S   H   A   V   F   T   I   V   F   T   Q
```

FIGURE 1B

```
          767         776         785         794         803         812
CGC TGC CAT GAC CAG CTC ACG GGG CTG GAC TCG GAG AAG GTC AGT AAG ATC AGT
R   C   H   D   Q   L   T   G   L   D   S   E   K   V   S   K   I   S

          821         830         839         848         857         866
TTG GTG GAC CTT GCT GGG AGT GAG CGA GCC GAC TCC TCA GGG GCC CGG GGC ATG
L   V   D   L   A   G   S   E   R   A   D   S   S   G   A   R   G   M

          875         884         893         902         911         920
GGC CTG AAG GAA GGA GCC AAC ATC AAT AAG TCC CTG ACT ACA CTA GGG AAA GTG
G   L   K   E   G   A   N   I   N   K   S   L   T   T   L   G   K   V

          929         938         947         956         965         974
ATC TCG GCC CTT GCA GAT ATG CAA TCA AAG AAG CGA AAG TCG GAT TTT ATC CCC
I   S   A   L   A   D   M   Q   S   K   K   R   K   S   D   F   I   P

          983         992        1001        1010        1019        1028
TAC AGG GAC TCT GTG CTC ACC TGG CTG CTC AAG GAA AAT TTG GGG GGG AAC TCA
Y   R   D   S   V   L   T   W   L   L   K   E   N   L   G   G   N   S

         1037        1046        1055        1064        1073        1082
CGC ACA GCC ATG ATT GCA GCC CTG AGC CCT GCT GAC ATC AAT TAC GAG GAG ACT
R   T   A   M   I   A   A   L   S   P   A   D   I   N   Y   E   E   T

         1091        1100        1109        1118        1127        1136
CTC AGC ACC CTC AGG TAT GCT GAC CGC ACC AAG CAA ATC CGC TGC AAT GCC ATC
L   S   T   L   R   Y   A   D   R   T   K   Q   I   R   C   N   A   I
```

FIGURE 1C

```
      1145        1154        1163        1172        1181        1190
ATC AAC GAG GAC CCT AAT GCC CGG CTG ATT AGA GAG CTG CAG GAG GAA GTA GCC
I   N   E   D   P   N   A   R   L   I   R   E   L   Q   E   E   V   A

      1199        1208        1217        1226        1235        1244
CGG CTG CGG GAA CTG CTG ATG GCT CAG GGA CTG TCA GCC TCT GCT CTG GAA GGC
R   L   R   E   L   L   M   A   Q   G   L   S   A   S   A   L   E   G

      1253        1262        1271        1280        1289        1298
CTG AAG ACG GAA GAA GGG AGT GTC AGA GGC GCC CTG CCA GCT GTG TCA TCT CCC
L   K   T   E   E   G   S   V   R   G   A   L   P   A   V   S   S   P

      1307        1316        1325        1334        1343        1352
CCA GCT CCA GTT TCA CCC TCA TCA CCC ACC ACA CAT AAT GGG GAG CTG GAG CCG
P   A   P   V   S   P   S   S   P   T   T   H   N   G   E   L   E   P

      1361        1370        1379        1388        1397        1406
TCA TTC TCC CCC AAC ACG GAG TCC CAG ATT GGG CCT GAG GAA GCC ATG GAG AGG
S   F   S   P   N   T   E   S   Q   I   G   P   E   E   A   M   E   R

      1415        1424        1433        1442        1451        1460
CTG CAG GAG ACA GAG AAG ATT ATA GCT GAG CTG AAC GAG ACA TGG GAG GAG AAG
L   Q   E   T   E   K   I   I   A   E   L   N   E   T   W   E   E   K

      1469        1478        1487        1496        1505        1514
CTA CGC AAG ACA GAA GCC CTG AGG ATG GAG AGA GAA GCA TTG CTG GCT GAG ATG
L   R   K   T   E   A   L   R   M   E   R   E   A   L   L   A   E   M
```

FIGURE 1D

```
      1523        1532        1541        1550        1559        1568
GGG GTG GCC GTC CGG GAG GAT GGG GGA ACT GTG GGC GTC TTC TCT CCA AAG AAG
G   V   A   V   R   E   D   G   G   T   V   G   V   F   S   P   K   K

      1577        1586        1595        1604        1613        1622
ACT CCC CAC CTG GTG AAC CTG AAC GAA GAC CCT CTG ATG TCT GAG TGT CTG CTC
T   P   H   L   V   N   L   N   E   D   P   L   M   S   E   C   L   L

      1631        1640        1649        1658        1667        1676
TAC CAC ATC AAA GAT GGC GTC ACC AGG GTC GGC CAA GTA GAT ATG GAC ATC AAG
Y   H   I   K   D   G   V   T   R   V   G   Q   V   D   M   D   I   K

      1685        1694        1703        1712        1721        1730
CTG ACC GGA CAG TTC ATT CGG GAG CAA CAC TGT CTG TTC CGG AGC ATC CCC CAG
L   T   G   Q   F   I   R   E   Q   H   C   L   F   R   S   I   P   Q

      1739        1748        1757        1766        1775        1784
CCA GAT GGA GAA GTG GTG GTC ACT CTG GAG CCT TGT GAA GGA GCT GAG ACA TAT
P   D   G   E   V   V   V   T   L   E   P   C   E   G   A   E   T   Y

      1793        1802        1811        1820        1829        1838
GTG AAT GGG AAG CTT GTG ACG GAG CCG CTG GTG CTG AAG TCA GGG AAT AGG ATT
V   N   G   K   L   V   T   E   P   L   V   L   K   S   G   N   R   I

      1847        1856        1865        1874        1883        1892
GTG ATG GGC AAG AAC CAC GTT TTC CGC TTC AAC CAC CCG GAG CAG GCA AGG CTG
V   M   G   K   N   H   V   F   R   F   N   H   P   E   Q   A   R   L
```

FIGURE 1E

```
      1901        1910        1919        1928        1937        1946
GAA CGG GAA CGA GGG GTC CCC CCA CCC CCA GGA CCG CCC TCT GAG CCA GTC GAC
E   R   E   R   G   V   P   P   P   P   G   P   P   S   E   P   V   D

      1955        1964        1973        1982        1991        2000
TGG AAC TTT GCC CAG AAG GAA CTG CTG GAG CAG CAA GGC ATC GAC ATA AAG CTG
W   N   F   A   Q   K   E   L   L   E   Q   Q   G   I   D   I   K   L

      2009        2018        2027        2036        2045        2054
GAA ATG GAG AAG AGG CTG CAG GAT CTG GAG AAT CAG TAC CGG AAA GAA AAG GAA
E   M   E   K   R   L   Q   D   L   E   N   Q   Y   R   K   E   K   E

      2063        2072        2081        2090        2099        2108
GAA GCC GAT CTT CTG CTG GAG CAG CAG CGA CTG TAT GCA GAC TCG GAC AGC GGG
E   A   D   L   L   L   E   Q   Q   R   L   Y   A   D   S   D   S   G

      2117        2126        2135        2144        2153        2162
GAT GAC TCT GAC AAG CGC TCT TGT GAA GAG AGC TGG AGG CTC ATC ACC TCC TTG
D   D   S   D   K   R   S   C   E   E   S   W   R   L   I   T   S   L

      2171        2180        2189        2198        2207        2216
CGG GAG CAG CTG CCG CCC ACC ACG GTC CAG ACC ATT GTC AAA CGC TGT GGT CTG
R   E   Q   L   P   P   T   T   V   Q   T   I   V   K   R   C   G   L

      2225        2234        2243        2252        2261        2270
CCC AGC AGT GGC AAG CGC AGG GCC CCT CGC AGG GTT TAT CAG ATC CCC CAG CGG
P   S   S   G   K   R   R   A   P   R   R   V   Y   Q   I   P   Q   R
```

FIGURE 1F

```
        2279          2288          2297          2306          2315          2324
CGC AGG CTG CAG GGC AAA GAC CCC CGC TGG GCC ACC ATG GCT GAC CTG AAG ATG
R   R   L   Q   G   K   D   P   R   W   A   T   M   A   D   L   K   M

        2333          2342          2351          2360          2369          2378
CAG GCG GTG AAG GAG ATC TGC TAC GAG GTG GCC CTG GCT GAC TTC CGC CAC GGG
Q   A   V   K   E   I   C   Y   E   V   A   L   A   D   F   R   H   G

        2387          2396          2405          2414          2423          2432
CGG GCT GAG ATT GAG GCC CTG GCC GCC CTC AAG ATG CGG GAG CTG TGT CGC ACC
R   A   E   I   E   A   L   A   A   L   K   M   R   E   L   C   R   T

        2441          2450          2459          2468          2477          2486
TAT GGC AAG CCA GAC GGC CCC GGA GAC GCC TGG AGG GCT GTG GCC CGG GAT GTC
Y   G   K   P   D   G   P   G   D   A   W   R   A   V   A   R   D   V

        2495          2504          2513          2522          2531          2540
TGG GAC ACT GTA GGC GAG GAG GAA GGA GGT GGA GCT GGC AGT GGT GGT GGC AGT
W   D   T   V   G   E   E   E   G   G   G   A   G   S   G   G   G   S

        2549          2558          2567          2576          2585          2594
GAG GAG GGA GCC CGA GGG GCG GAG GTG GAG GAC CTC CGG GCC CAC ATC GAC AAG
E   E   G   A   R   G   A   E   V   E   D   L   R   A   H   I   D   K

        2603          2612          2621          2630          2639          2648
CTG ACG GGG ATT CTG CAG GAG GTG AAG CTG CAG AAC AGC AGC AAG GAC CGG GAG
L   T   G   I   L   Q   E   V   K   L   Q   N   S   S   K   D   R   E
```

FIGURE 1G

PCT/US99/22495

```
        2657          2666          2675          2684          2693          2702
CTG CAG GCC CTG CGG GAC CGC ATG CTC CGC ATG GAG AGG GTC ATC CCC CTG GCC
L   Q   A   L   R   D   R   M   L   R   M   E   R   V   I   P   L   A

        2711          2720          2729          2738          2747          2756
CAG GAT CAT GAG GAT GAG AAT GAA GAA GGT GGT GAG GTC CCC TGG GCC CCG CCT
Q   D   H   E   D   E   N   E   E   G   G   E   V   P   W   A   P   P

        2765          2774          2783          2792          2801          2810
GAA GGA TCA GAG GCA GCA GAG GAG GCA GCC CCC AGT GAC CGC ATG CCG TCA GCC
E   G   S   E   A   A   E   E   A   A   P   S   D   R   M   P   S   A

        2819          2828          2837          2846          2855          2864
CGG CCC CCC TCG CCG CCA CTG TCA AGC TGG GAG CGG GTG TCA CGG CTC ATG GAG
R   P   P   S   P   P   L   S   S   W   E   R   V   S   R   L   M   E

        2873          2882          2891          2900          2909          2918
GAG GAC CCT GCC TTC CGT CGT GGT CGT CTT CGC TGG CTC AAG CAG GAG CAG CTA
E   D   P   A   F   R   R   G   R   L   R   W   L   K   Q   E   Q   L

        2927          2936          2945          2954          2963          2972
CGG CTG CAG GGA CTG CAG GGC TCT GGG GGC CGG GGC GGG GGG CTG CGC AGG CCC
R   L   Q   G   L   Q   G   S   G   G   R   G   G   G   L   R   R   P

        2981          2990          2999          3008          3017          3026
CCA GCC CGC TTT GTG CCC CCT CAC GAC TGC AAG CTA CGC TTC CCC TTC AAG AGC
P   A   R   F   V   P   P   H   D   C   K   L   R   F   P   F   K   S
```

FIGURE 1H

```
      3035        3044        3053        3062        3071        3080
AAC CCC CAG CAC CGG GAG TCT TGG CCA GGG ATG GGG AGC GGG GAG GCT CCA ACT
N   P   Q   H   R   E   S   W   P   G   M   G   S   G   E   A   P   T

      3089        3098        3107        3116        3125        3134
CCG CTC CAA CCC CCT GAG GAG GTC ACT CCC CAT CCA GCC ACC CCT GCC CGC CGG
P   L   Q   P   P   E   E   V   T   P   H   P   A   T   P   A   R   R

      3143        3152        3161        3170        3179        3188
CCT CCG AGT CCC CGA AGG TCC CAC CAT CCC CGC AGG AAC TCC CTG GAT GGA GGG
P   P   S   P   R   R   S   H   H   P   R   R   N   S   L   D   G   G

      3197        3206        3215        3224        3233        3242
GGC CGA TCC CGG GGA GCG GGT TCT GCA CAG CCT GAA CCC CAG CAC TTC CAG CCC
G   R   S   R   G   A   G   S   A   Q   P   E   P   Q   H   F   Q   P

      3251        3260        3269        3278        3287        3296
AAA AAG CAC AAC TCT TAT CCC CAG CCA CCC CAA CCC TAC CCA GCC CAG CGG CCC
K   K   H   N   S   Y   P   Q   P   P   Q   P   Y   P   A   Q   R   P

      3305        3314        3323        3332        3341        3350
CCA GGG CCC CGC TAC CCC CCA TAC ACT ACT CCC CCA CGA ATG AGA CGG CAG CGT
P   G   P   R   Y   P   P   Y   T   T   P   P   R   M   R   R   Q   R

      3359        3368        3377        3386        3395        3404
TCT GCC CCT GAC CTC AAG GAG AGT GGG GCA GCT GTG TGA GTC CCA CAT CCT GGG
S   A   P   D   L   K   E   S   G   A   A   V
```

FIGURE 11

```
       3413        3422        3431        3440        3449        3458
CAG AGG GCC TGG TGG GGC CCC TTG CTA GGA GAA GGG AAG ACG CCC GAG ACG CTG

       3467        3476        3485        3494        3503        3512
CTT CCC CAG AAG TGC TGG GGC AGG GAG GCC CAG GAG ATG AGA GAG AAG GTC CGA

       3521        3530        3539        3548        3557        3566
GTA GGT GAT AGA AGA CAA GGG GGA GAC CGA GCC GGA GGC TGA GGA AAG GAA GAG

       3575        3584        3593        3602        3611        3620
GGC ACG GAG TTG CCA GGA GCA AAC CAA AGT GAA GAG AGA GAT AGG AAG CTG CCT

       3629        3638        3647        3656        3665        3674
CGG GGC CAC CCC TTG CAA AGG GGG TGT GTC CCA CAA ACG CTG CTA TGG GTG GGG

       3683        3692        3701        3710        3719        3728
TGG GGG GCT GGG GTG CTG CGT AGC CAG TGT TTG ACT TTC TTT TCA AGT GGG GGA

       3737        3746        3755        3764        3773        3782
AAG TGG GAG AGG ACT GAG AGT GAG GCA AGT TCT CCC CAG CCC CTG TCC GTC TGT

       3791        3800        3809        3818        3827        3836
CTG TCT CTG TCT GTG GTG GTT TCT GTT TCT TGG GAG GCA TGG TAG GAT CAT AAG

       3845        3854        3863        3872        3881        3890
TCA TTC CCC TCC CCT TCC AGG CCT CCT GCT ATA TTT GGG GGA CCT GAC TGG TTT

       3899        3908        3917        3926
GGC TGG AGT CCC GAT GAG GAT GTG GCC CTT ACT ATA GGT A 3'
```

FIGURE 1J

Library	Library Description	Abundance	Percent Abundance
MUSCNOT11	muscle, arm, ALS, 74F	5	0.1314
BRAYDIT01	brain, hypothalamus, Huntington's disease, aw/CVA, 57M	3	0.0886
PONSAZT01	brain, pons, AD, 74M	5	0.0881

FIGURE 2

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
TANG, Y. Tom
CORLEY, Neil C.
GUEGLER, Karl J.
PATTERSON, Chandra

<120> KINESIN-LIKE MOTOR PROTEIN

<130> PF-0593 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/162,373
<151> 1998-09-28

<160> 2

<170> PERL Program

<210> 1
<211> 1103
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1281811CD1

<400> 1

Met	Ala	Gly	Ala	Ser	Val	Lys	Val	Ala	Val	Arg	Val	Arg	Pro	Phe			
1				5					10					15			
Asn	Ala	Arg	Glu	Thr	Ser	Gln	Asp	Ala	Lys	Cys	Val	Val	Ser	Met			
				20					25					30			
Gln	Gly	Asn	Thr	Thr	Ser	Ile	Ile	Asn	Pro	Lys	Gln	Ser	Lys	Asp			
				35					40					45			
Ala	Pro	Lys	Ser	Phe	Thr	Phe	Asp	Tyr	Ser	Tyr	Trp	Ser	His	Thr			
				50					55					60			
Ser	Thr	Glu	Asp	Pro	Gln	Phe	Ala	Ser	Gln	Gln	Gln	Val	Tyr	Arg			
				65					70					75			
Asp	Ile	Gly	Glu	Glu	Met	Leu	Leu	His	Ala	Phe	Glu	Gly	Tyr	Asn			
				80					85					90			
Val	Cys	Ile	Phe	Ala	Tyr	Gly	Gln	Thr	Gly	Ala	Gly	Lys	Ser	Tyr			
				95					100					105			
Thr	Met	Met	Gly	Arg	Gln	Glu	Pro	Gly	Gln	Gln	Gly	Ile	Val	Pro			
				110					115					120			
Gln	Leu	Cys	Glu	Asp	Leu	Phe	Ser	Arg	Val	Ser	Glu	Asn	Gln	Ser			
				125					130					135			
Ala	Gln	Leu	Ser	Tyr	Ser	Val	Glu	Val	Ser	Tyr	Met	Glu	Ile	Tyr			
				140					145					150			
Cys	Glu	Arg	Val	Arg	Asp	Leu	Leu	Asn	Pro	Lys	Ser	Arg	Gly	Ser			
				155					160					165			
Leu	Arg	Val	Arg	Glu	His	Pro	Ile	Leu	Gly	Pro	Tyr	Val	Gln	Asp			
				170					175					180			

Leu Ser Lys Leu Ala Val Thr Ser Tyr Ala Asp Ile Ala Asp Leu
185 190 195
Met Asp Cys Gly Asn Lys Ala Arg Thr Val Ala Ala Thr Asn Met
200 205 210
Asn Glu Thr Ser Ser Arg Ser His Ala Val Phe Thr Ile Val Phe
215 220 225
Thr Gln Arg Cys His Asp Gln Leu Thr Gly Leu Asp Ser Glu Lys
230 235 240
Val Ser Lys Ile Ser Leu Val Asp Leu Ala Gly Ser Glu Arg Ala
245 250 255
Asp Ser Ser Gly Ala Arg Gly Met Gly Leu Lys Glu Gly Ala Asn
260 265 270
Ile Asn Lys Ser Leu Thr Thr Leu Gly Lys Val Ile Ser Ala Leu
275 280 285
Ala Asp Met Gln Ser Lys Lys Arg Lys Ser Asp Phe Ile Pro Tyr
290 295 300
Arg Asp Ser Val Leu Thr Trp Leu Leu Lys Glu Asn Leu Gly Gly
305 310 315
Asn Ser Arg Thr Ala Met Ile Ala Ala Leu Ser Pro Ala Asp Ile
320 325 330
Asn Tyr Glu Glu Thr Leu Ser Thr Leu Arg Tyr Ala Asp Arg Thr
335 340 345
Lys Gln Ile Arg Cys Asn Ala Ile Ile Asn Glu Asp Pro Asn Ala
350 355 360
Arg Leu Ile Arg Glu Leu Gln Glu Glu Val Ala Arg Leu Arg Glu
365 370 375
Leu Leu Met Ala Gln Gly Leu Ser Ala Ser Ala Leu Glu Gly Leu
380 385 390
Lys Thr Glu Glu Gly Ser Val Arg Gly Ala Leu Pro Ala Val Ser
395 400 405
Ser Pro Pro Ala Pro Val Ser Pro Ser Ser Pro Thr Thr His Asn
410 415 420
Gly Glu Leu Glu Pro Ser Phe Ser Pro Asn Thr Glu Ser Gln Ile
425 430 435
Gly Pro Glu Glu Ala Met Glu Arg Leu Gln Glu Thr Glu Lys Ile
440 445 450
Ile Ala Glu Leu Asn Glu Thr Trp Glu Glu Lys Leu Arg Lys Thr
455 460 465
Glu Ala Leu Arg Met Glu Arg Glu Ala Leu Leu Ala Glu Met Gly
470 475 480
Val Ala Val Arg Glu Asp Gly Gly Thr Val Gly Val Phe Ser Pro
485 490 495
Lys Lys Thr Pro His Leu Val Asn Leu Asn Glu Asp Pro Leu Met
500 505 510
Ser Glu Cys Leu Leu Tyr His Ile Lys Asp Gly Val Thr Arg Val
515 520 525
Gly Gln Val Asp Met Asp Ile Lys Leu Thr Gly Gln Phe Ile Arg
530 535 540
Glu Gln His Cys Leu Phe Arg Ser Ile Pro Gln Pro Asp Gly Glu
545 550 555
Val Val Val Thr Leu Glu Pro Cys Glu Gly Ala Glu Thr Tyr Val
560 565 570
Asn Gly Lys Leu Val Thr Glu Pro Leu Val Leu Lys Ser Gly Asn
575 580 585
Arg Ile Val Met Gly Lys Asn His Val Phe Arg Phe Asn His Pro
590 595 600
Glu Gln Ala Arg Leu Glu Arg Glu Arg Gly Val Pro Pro Pro Pro

```
                605                 610                 615
Gly Pro Pro Ser Glu Pro Val Asp Trp Asn Phe Ala Gln Lys Glu
                620                 625                 630
Leu Leu Glu Gln Gln Gly Ile Asp Ile Lys Leu Glu Met Glu Lys
                635                 640                 645
Arg Leu Gln Asp Leu Glu Asn Gln Tyr Arg Lys Glu Lys Glu Glu
                650                 655                 660
Ala Asp Leu Leu Leu Glu Gln Gln Arg Leu Tyr Ala Asp Ser Asp
                665                 670                 675
Ser Gly Asp Asp Ser Asp Lys Arg Ser Cys Glu Glu Ser Trp Arg
                680                 685                 690
Leu Ile Thr Ser Leu Arg Glu Gln Leu Pro Pro Thr Thr Val Gln
                695                 700                 705
Thr Ile Val Lys Arg Cys Gly Leu Pro Ser Ser Gly Lys Arg Arg
                710                 715                 720
Ala Pro Arg Arg Val Tyr Gln Ile Pro Gln Arg Arg Arg Leu Gln
                725                 730                 735
Gly Lys Asp Pro Arg Trp Ala Thr Met Ala Asp Leu Lys Met Gln
                740                 745                 750
Ala Val Lys Glu Ile Cys Tyr Glu Val Ala Leu Ala Asp Phe Arg
                755                 760                 765
His Gly Arg Ala Glu Ile Glu Ala Leu Ala Ala Leu Lys Met Arg
                770                 775                 780
Glu Leu Cys Arg Thr Tyr Gly Lys Pro Asp Gly Pro Gly Asp Ala
                785                 790                 795
Trp Arg Ala Val Ala Arg Asp Val Trp Asp Thr Val Gly Glu Glu
                800                 805                 810
Glu Gly Gly Gly Ala Gly Ser Gly Gly Gly Ser Glu Glu Gly Ala
                815                 820                 825
Arg Gly Ala Glu Val Glu Asp Leu Arg Ala His Ile Asp Lys Leu
                830                 835                 840
Thr Gly Ile Leu Gln Glu Val Lys Leu Gln Asn Ser Ser Lys Asp
                845                 850                 855
Arg Glu Leu Gln Ala Leu Arg Asp Arg Met Leu Arg Met Glu Arg
                860                 865                 870
Val Ile Pro Leu Ala Gln Asp His Glu Asp Glu Asn Glu Glu Gly
                875                 880                 885
Gly Glu Val Pro Trp Ala Pro Pro Glu Gly Ser Glu Ala Ala Glu
                890                 895                 900
Glu Ala Ala Pro Ser Asp Arg Met Pro Ser Ala Arg Pro Pro Ser
                905                 910                 915
Pro Pro Leu Ser Ser Trp Glu Arg Val Ser Arg Leu Met Glu Glu
                920                 925                 930
Asp Pro Ala Phe Arg Arg Gly Arg Leu Arg Trp Leu Lys Gln Glu
                935                 940                 945
Gln Leu Arg Leu Gln Gly Leu Gln Gly Ser Gly Gly Arg Gly Gly
                950                 955                 960
Gly Leu Arg Arg Pro Pro Ala Arg Phe Val Pro Pro His Asp Cys
                965                 970                 975
Lys Leu Arg Phe Pro Phe Lys Ser Asn Pro Gln His Arg Glu Ser
                980                 985                 990
Trp Pro Gly Met Gly Ser Gly Glu Ala Pro Thr Pro Leu Gln Pro
                995                1000                1005
Pro Glu Glu Val Thr Pro His Pro Ala Thr Pro Ala Arg Arg Pro
               1010                1015                1020
Pro Ser Pro Arg Arg Ser His His Pro Arg Arg Asn Ser Leu Asp
               1025                1030                1035
```

Gly Gly Gly Arg Ser Arg Gly Ala Gly Ser Ala Gln Pro Glu Pro
1040 1045 1050
Gln His Phe Gln Pro Lys Lys His Asn Ser Tyr Pro Gln Pro Pro
1055 1060 1065
Gln Pro Tyr Pro Ala Gln Arg Pro Pro Gly Pro Arg Tyr Pro Pro
1070 1075 1080
Tyr Thr Thr Pro Pro Arg Met Arg Arg Gln Arg Ser Ala Pro Asp
1085 1090 1095
Leu Lys Glu Ser Gly Ala Ala Val
1100

<210> 2
<211> 3930
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1281811CB1

<400> 2
gtggcagcca gaactgatac agccccccctg gtctggggcc aggacgccag ctgaggaggg 60
caggagtgtc tggagctatg gctggtgcct cggtgaaagt ggcagtgagg gttcggccct 120
ttaacgcccg tgagaccagc caggatgcca agtgtgtggt cagcatgcag ggcaacacca 180
cctccatcat caatcctaaa cagagcaagg atgcccccaa aagcttcacc tttgactact 240
cctactggtc acacacttcg acggaggacc cccagtttgc atctcagcag caagtgtatc 300
gggacattgg agaagagatg ctgctccacg cctttgaagg ctacaacgtg tgcatctttg 360
cctatgggca gaccggggct gggaaatcct ataccatgat ggggcgacag gagccagggc 420
agcagggcat cgtgccccag ctctgtgagg acctcttctc tcgcgttagt gagaaccaga 480
gtgctcagct atcctactct gtggaggtga gctatatgga gatctactgt gagcgggtac 540
gagacctctt gaaccccaag agtcggggtt ctctgcgggt ccgggagcac cccatcctgg 600
gcccgtacgt gcaggacctg tccaaattgg ctgtgacctc ctacgcagac attgctgacc 660
tcatggactg tggaaataaa gcacggactg tggctgccac caacatgaat gagaccagca 720
gccgttccca tgccgtcttt accatcgtct tcacacagcg ctgccatgac cagctcacgg 780
ggctggactc ggagaaggtc agtaagatca gtttggtgga ccttgctggg agtgagcgag 840
ccgactcctc aggggcccgg ggcatgggcc tgaaggaagg agccaacatc aataagtccc 900
tgactacact agggaaagtg atctcggccc ttgcagatat gcaatcaaag aagcgaaagt 960
cggattttat cccctacagg gactctgtgc tcacctggct gctcaaggaa aatttggggg 1020
ggaactcacg cacagccatg attgcagccc tgagccctgc tgacatcaat tacgaggaga 1080
ctctcagcac cctcaggtat gctgaccgca ccaagcaaat ccgctgcaat gccatcatca 1140
acgaggaccc taatgcccgg ctgattagag agctgcagga ggaagtagcc cggctgcggg 1200
aactgctgat ggctcaggga ctgtcagcct ctgctctgga aggcctgaag acggaagaag 1260
ggagtgtcag aggcgccctg ccagctgtgt catctccccc agctccagtt tcaccctcat 1320
cacccaccac acataatggg gagctggagc cgtcattctc ccccaacacg gagtcccaga 1380
ttgggcctga ggaagccatg gagaggctgc aggagacaga gaagattata gctgagctga 1440
acgagacatg ggaggagaag ctacgcaaga cagaagccct gaggatggag agagaagcat 1500
tgctggctga gatgggggtg gccgtccggg aggatggggg aactgtgggc gtcttctctc 1560
caaagaagac tccccacctg gtgaacctga acgaagaccc tctgatgtct gagtgtctgc 1620
tctaccacat caaagatggc gtcaccaggg tcggccaagt agatatggac atcaagctga 1680
ccggacagtt cattcgggag caacactgtc tgttccggag catcccccag ccagatggag 1740
aagtggtggt cactctggag ccttgtgaag gagctgagac atatgtgaat gggaagcttg 1800
tgacggagcc gctggtgctg aagtcaggga ataggattgt gatgggcaag aaccacgttt 1860
tccgcttcaa ccacccggag caggcaaggc tggaacggga acgaggggtc cccccacccc 1920
caggaccgcc ctctgagcca gtcgactgga actttgccca gaaggaactg ctggagcagc 1980
aaggcatcga cataaagctg gaaatggaga agaggctgca ggatctggag aatcagtacc 2040
ggaaagaaaa ggaagaagcc gatcttctgc tggagcagca gcgactgtat gcagactcgg 2100

acagcgggga tgactctgac aagcgctctt gtgaagagag ctggaggctc atcacctcct 2160
tgcgggagca gctgccgccc accacggtcc agaccattgt caaacgctgt ggtctgccca 2220
gcagtggcaa gcgcagggcc cctcgcaggg tttatcagat cccccagcgg cgcaggctgc 2280
agggcaaaga cccccgctgg gccaccatgg ctgacctgaa gatgcaggcg gtgaaggaga 2340
tctgctacga ggtggccctg gctgacttcc gccacgggcg ggctgagatt gaggccctgg 2400
ccgccctcaa gatgcgggag ctgtgtcgca cctatggcaa gccagacggc cccggagacg 2460
cctggagggc tgtggcccgg gatgtctggg acactgtagg cgaggaggaa ggaggtggag 2520
ctggcagtgg tggtggcagt gaggagggag cccgaggggc ggaggtggag gacctccggg 2580
cccacatcga caagctgacg gggattctgc aggaggtgaa gctgcagaac agcagcaagg 2640
accgggagct gcaggccctg cgggaccgca tgctccgcat ggagagggtc atccccctgg 2700
cccaggatca tgaggatgag aatgaagaag gtggtgaggt cccctgggcc ccgcctgaag 2760
gatcagaggc agcagaggag gcagccccca gtgaccgcat gccgtcagcc cggcccccct 2820
cgccgccact gtcaagctgg gagcgggtgt cacggctcat ggaggaggac cctgccttcc 2880
gtcgtggtcg tcttcgctgg ctcaagcagg agcagctacg gctgcaggga ctgcagggct 2940
ctgggggccg gggcgggggg ctgcgcaggc ccccagcccg ctttgtgccc cctcacgact 3000
gcaagctacg cttccccttc aagagcaacc cccagcaccg ggagtcttgg ccagggatgg 3060
ggagcgggga ggctccaact ccgctccaac cccctgagga ggtcactccc catccagcca 3120
cccctgcccg ccggcctccg agtccccgaa ggtcccacca tccccgcagg aactccctgg 3180
atggaggggg ccgatcccgg ggagcgggtt ctgcacagcc tgaaccccag cacttccagc 3240
ccaaaaagca caactcttat ccccagccac cccaacccta cccagcccag cggcccccag 3300
ggccccgcta ccccccatac actactcccc cacgaatgag acggcagcgt tctgcccctg 3360
acctcaagga gagtggggca gctgtgtgag tcccacatcc tgggcagagg gcctggtggg 3420
gccccttgct aggagaaggg aagacgcccg agacgctgct tccccagaag tgctggggca 3480
gggaggccca ggagatgaga gagaaggtcc gagtaggtga tagaagacaa gggggagacc 3540
gagccggagg ctgaggaaag gaagagggca cggagttgcc aggagcaaac caaagtgaag 3600
agagagatag gaagctgcct cggggccacc ccttgcaaag ggggtgtgtc ccacaaacgc 3660
tgctatgggt ggggtggggg gctggggtgc tgcgtagcca gtgtttgact ttcttttcaa 3720
gtgggggaaa gtgggagagg actgagagtg aggcaagttc tccccagccc ctgtccgtct 3780
gtctgtctct gtctgtggtg gtttctgttt cttgggaggc atggtaggat cataagtcat 3840
tcccctcccc ttccaggcct cctgctatat ttgggggacc tgactggttt ggctggagtc 3900
ccgatgagga tgtggccctt actataggta 3930

INTERNATIONAL SEARCH REPORT

Intern. nal Application No
PCT/US 99/22495

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C12N5/10 C07K16/18 A61K38/17
C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K C12Q A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENEMBL 'Online! 8 January 1998 (1998-01-08) DORNER ET AL.: "Human kinesin-like motor protein KIF1c mRNA, complete cds." XP002125366 Accession U91329 -& DORNER,C. ET AL: "Characterization of KIF1C, a new kinesin-like protein involved in vesicle transport from the Golgi apparatus to the endoplasmic reticulum" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 32, 7 August 1998 (1998-08-07), pages 20267-20275, XP002125365 the whole document --- -/--	1-16,19

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 January 2000	25/01/2000
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer ALCONADA RODRIG.., A

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Intern al Application No
PCT/US 99/22495

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENEMBL 'Online! 15 July 1998 (1998-07-15) OHARA ET AL.: "Homo sapiens mRNA for KIAA0706 protein, complete cds." XP002125367 Accession AB014606 -& OHARA ET AL: "Prediction of the coding sequences of unidentified human genes. X. The complete coding sequences of 100 new cDNA clones from brain which can code for large proteins in vitro." DNA RESEARCH, vol. 5, 30 June 1998 (1998-06-30), pages 169-176, XP002089186 table 2	1-16,19

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/22495

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim 19 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 17,18,20
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
See FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17,18,20

Claims 17 and 18 refer to an agonist/activator and to an antagonist/inhibitor respectively, of the polypeptide, and claim 20 to a method of treatment of a subject with the antagonist of claim 18, without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambigous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT).
No search could be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the result to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.